Developing a logistics service quality scale

John T Mentzer, Daniel J Flint, John L Kent. **Journal of Business Logistics**. Oak Brook: 1999.Vol.20, Iss. 1; pg. 9, 24 pgs

» Jump to full text

Subjects:	Logistics, Measurement, Quality of service, Defense industry, Studies
Classification Codes	9190 United States, 9550 Public sector, 5160 Transportation management, 5320 Quality control, 9
Locations:	United States, US
Companies:	Defense Logistics Agency
Author(s):	John T Mentzer, Daniel J Flint, John L Kent
Article types:	Feature
Publication title:	Journal of Business Logistics. Oak Brook: 1999. Vol. 20, Iss. 1; pg. 9, 24 pgs
Source Type:	Periodical
ISSN/ISBN:	07353766
ProQuest document ID:	40452078
Text Word Count	6538
Article URL:	http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&res_dat=xri:pqd&rft_val_fmt=info:ofi/fmt:kev:mtx:journal&genre=article&rft_dat=xri:pqd:did=000

Abstract (Article Summary)

Building on past research and focus groups, a scale to measure logistics service quality was conceptualized, tested on a large sample, refined through confirmatory factor analyses, and validated within eight market segments of the Defense Logistics Agency. The resulting scale enhances the ability of managers and researchers to more precisely measure customer-perceived aspects of logistics services.

Full Text (6538 words)

In the pursuit of competitive advantage, it is increasingly important to identify the demands and values of current and potential customers.' Traditionally, logistics managers have done an excellent job of managing and moving inventory-the operational aspects of logistics. They often struggle to identify the perceptual impact of customer service activities associated with logistics-i.e., the marketing aspects of logistics.2

The preponderance of the marketing literature addressing customer service, or more specifically service quality, has been aimed at the end-use customer.3 Several authors, however, have attempted to expand the theoretical domain of service quality to a business-to-business context, specifically in the arena of logistics service quality. In particular, Bienstock, Mentzer, and Bird developed a valid, reliable scale of what they termed physical distribution service quality, or PDSQ, through surveying abroad range of purchasing managers.5 The process they followed was based upon the logistics research framework suggested by Mentzer and Kahn.6

It is the purpose of this paper to continue this expansion of the service quality domain into a logistics context. This research investigates a particular focal organization with multiple market segments to determine whether the general methodology used by Bienstock, Mentzer, and Bird results in a similarly valid, reliable scale of logistics service quality (LSQ). By doing so, the work begun by Brensinger and Lambert and Bienstock, Mentzer, and Bird in broad applications (i.e., across a number of firms) can be focused on the logistics customer service environment faced by one focal organization.7

To accomplish this purpose, the next two sections review the literature relevant to logistics service quality and service quality in general. This is followed by a description of a study undertaken to expand the application of a logistics service quality scale into a specific logistics context. This research is partially a response to recent calls for theoretically and methodologically rigorous, yet managerially relevant, research in logistics.8 Thus, in the last sections, the managerial and research implications of the development of this logistics service quality scale are discussed.

LOGISTICS SERVICE QUALITY

There are many definitions and descriptions of how logistics creates customer satisfaction. The most traditional are based on the creation of time and place utility.9 The so-called "Seven Rs" describe the attributes of the company's product/service offering that lead to utility creation through logistics service, i.e., part of a product's marketing offering is the company's ability to deliver the right amount of the right product at the right place at the right time in the right condition at the right price with the right information.'

This definition implies that part of the value of a product is created by logistics service. Examples of historical, operational measures of logistics customer service are percent of items in stock, percent of orders delivered on time, percent of delivered items undamaged, etc. (For a more complete list of these operational measures, see Mentzer, Gomes, and Krapfel"). These attributes are considered the "value" provided by the logistics service dimensions of availability, timeliness, and condition.'2

As the business environment has changed, the operations-based definitions of logistics service have evolved. The basic concept of utility creation became inadequate to fully express the value created by logistics. The idea of value has been broadened to include numerous value-added operational tasks, such as packaging, third-party inventory management, bar coding, and information systems.'3 The value-added concept expanded the traditional time and place utilities to include form utility,' but it was still an operations-based concept. LaLonde and Zinszer described customer service as possessing three components: (1) an activity to satisfy customers' needs, (2) performance measures to ensure customer satisfaction, and (3) a philosophy of firm-wide commitment.'5

However, these components are all focused on the provider firm, not on the customer. Similarly, other research has developed a framework for quantifying the value created by logistics operations.'6 Although this research incorporates internal and external customers, it is also focused predominately on provider firms-that is, how logistics executives can quantify the value they create for customers. We still must develop an instrument to measure customer perceptions of the value created for them by logistics services.

Regarding customer service in the logistics service context, Mentzer, Gomes, and Krapfel argue there are two elements in service delivery: marketing customer service (MCS) and physical distribution service (PDS). They recognize the complementary nature of the two elements to satisfy the customer and propose an integrative framework of customer service."7 This view is shared by others,18 and it is regarded as an intellectual base for integrating marketing and logistics activities.

The service quality approach is an attempt to understand customer satisfaction from the perspective of the differences between customer perceptions and actual customer service on various attributes.'9 Researchers have begun to examine whether the service quality model can be used to measure logistics service. Modifications have been made to the original service quality model by developing logistics attributes that fit into the previously customer-defined dimensions and by identifying additional gaps that could be applied to the logistics service context.20 These views of logistics service provide the building blocks to create a customer-based foundation for better definitions and measures of logistics service quality.

This use of customer-based definitions of LSQ brings physical distribution research, which traditionally has focused on more physically observable operational attributes, more in line with marketing, which has devoted more attention to understanding such unobservables as customers' perceived value. By recognizing, tapping into, and measuring

customer perceptions of LSQ service quality, logistics practitioners and researchers can add to the traditionally measured set of operational service attributes.

SERVICE QUALITY

A number of researchers have tried to empirically replicate the five dimensional structure (tangibles, responsiveness, empathy, reliability, and assurance) of the original Parasuraman, Zeithaml, and Berry service quality instrument, SERVQUAL.21 In developing it, Parasuraman, Zeithaml, and Berry followed a general procedure of qualitative research (interviews and focus groups) to develop the initial scale, followed by quantitative surveys to refine and verify the scale. These interviews and surveys included retail consumers of appliance repair or maintenance, retail banking, long-distance telephone service, securities brokers, and credit card services. Additional research has expanded the use of SERVQUAL to include retail consumers of health care, residential utilities, job placement, pest control, dry cleaning, financial services, and fast food services, with the resultant dimensions ranging from one to eight.22

Several researchers have argued for the addition of items and/or dimensions to SERVQUAL. Crosby defined quality as conformance to requirements and argued those requirements should be specifically defined to measure quality.3 In applying SERVQUAL to measure perceived quality of retail financial services, Brown, Churchill, and Peter noted the "omission of items we a priori thought would be critical to subjects' evaluation of...quality" (p. 138).24 Brensinger and Lambert, applying SERVQUAL to industrial purchasing of motor carrier transportation services, developed a fourfactor structure and recommended future research supplement SERVQUAL items with "service specific variables" (p. 289) to increase the validity of service quality measurement in an industrial service context.25

Bienstock, Mentzer, and Bird26 took note of these shortcomings in applying the concept of service quality to an industrial marketing context, and suggested a classification scheme based upon the work of Lovelock,n' Gronroos,28 and Parasuraman, Zeithaml, and Berry.29 Within this classification scheme, the consumer applications of SERVQUAL are in the context of people receiving intangible actions (services) that are not physically separated from the consumer. Bienstock, Mentzer, and Bird argue that business-to-business logistics services are offered in a context in which people are replaced with "things" (p. 34), and the customer and provider are physically separated. They maintain that the former is appropriate for the SERVQUAL "functional or process dimensions" (p. 33), but the latter logistics service context is composed more of "technical or outcome dimensions" (p. 34). They conclude that an "alternative conceptualization" is necessary for logistics service quality.30 As did Parasuraman, Zeithaml, and Berry,3' Bienstock, Mentzer, and Bird followed a methodology of a qualitative phase to develop the scale, followed by a quantitative survey to refine and verify it.

RESEARCH PURPOSE AND SETTING

The purpose of this research is to continue the expansion of the service quality domain into a logistics context. This research investigates a particular focal organization with multiple market segments in order to determine whether the general methodology used by Bienstock, Mentzer, and Bird results in a similarly valid, reliable scale of logistics service quality, LSQ. Such an expansion takes more of a customer-based approach to logistics than the traditional operational-based approaches to logistics customer service. Thus, it enhances the knowledge logistics managers have about their customers. Since this enhancement provides additional information to logistics managers on what truly motivates their customers to use one supplier over another, it has considerable value as competitive advantage input to the strategic planning goal of achieving "logistics leverage."32

The organization selected, the Defense Logistics Agency (DLA), performs logistical functions for the military services, Department of Defense agencies, and allied foreign governments. Currently, DLA manages over $22 billion in sales and over 4.5 million items. Thus, DLA acts in the role of a supplier in an industrial marketing context. At the time of this study, DLA's environment was becoming increasingly competitive as customers were given expanded choices in logistics service providers. DLA commissioned a research team to help improve the measurement and management of its logistics customer service in order to compete more effectively.

METHODOLOGY

The methodology closely followed that of Parasuraman, Zeithaml, and Berry33 and Bienstock, Mentzer, and Bird34 to develop the Logistics Service Quality (LSQ) scale. The first step was a qualitative effort to understand the LSQ needs of DLA's customers. From this qualitative effort, dimensions of LSQ for this particular industrial marketing

setting were developed, and survey instrument items were derived. The instrument was then tested with a broader array of DLA customers.

QUALITATIVE METHODOLOGY

This step involved thirteen focus group sessions, each lasting approximately two hours. The first author acted as the moderator for all the sessions. Since the goal was to have DLAcustomers express their own unfiltered views through the interaction of the group, the moderator provided only general topics to start the meeting. This unstructured technique is appropriate when the objective is to identify underlying themes."

The participants and topics varied across the focus groups, but all focused on assessments of logistics services. Each participant was identified with a particular market segment of DLA: medical supplies, fuels, electronics, clothing/textiles, construction, industrial supplies, subsistence, and general supplies. The individuals in each focus group were a mixture of DLA customers for that particular product category (for example, the fuels session included civilian and military personnel with fuel logistics responsibilities). The general topics covered four basic areas: (1) the nature of the participants' work in relation to DLA; (2) evaluation of the working relationship with DLA; (3) assessment of DLA performance; and (4) perceptions of what DLA does well or poorly.

Each focus group session was recorded on videotape for later analysis, and extensive notes were taken by the first author. The videotapes were reviewed by the entire research team. The researchers individually identified specific LSQ dimensions raised in the sessions, then met to discuss them and resolve any differences. From this stage emerged an initial set of eight LSQ dimensions specific to this industrial marketing setting: information quality, ordering procedures, timeliness, order accuracy, order quality, order condition, order discrepancy handling, and personal contact quality (Order accuracy here reflects product availability). From this analysis, specific items for each dimension were derived for the survey questionnaire. The intent was to develop items that would constitute a single scale for use across multiple market segments. Although different segments might place varying weight on each dimension (such as, timeliness over order accuracy in one segment and vice versa in another segment), the items that comprised all the dimensions of LSQ were intended to apply across all segments.

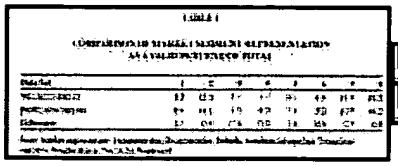
QUANTITATIVE METHODOLOGY

DLA personnel distributed the questionnaire to 16,920 DLA customers. The 5,531 usable questionnaires returned represented a response rate of 32.7%. The response sample represented all eight DLA market segments. Demographic variations were similar to the population at large, lending support to a conclusion of lack of nonresponse bias. All items used a standard 5-point Likert-type scale ranging from strongly disagree (1) to strongly agree (5).

OVERVIEW OF SCALE DEVELOPMENT

The 5,531 respondents were divided into ten separate data sets, one for scale purification, eight for scale validation within each of the eight market segments, and one for respondents who failed to indicate the segment to which they belonged. The latter data set was excluded from all data analysis. The purification set consisted of a random sample of 550 from the entire data set. Table I presents a comparison of the percent represented by each market segment in both the total data set and the purification data set. All eight market segments were represented in the randomly selected purification data set, with each segment representing a nearly equivalent proportion (i.e., largest difference was 1.6 percent) compared to the validation data sets.

Scale purification was conducted with the first data set through confirmatory factor analyses (CFA), eliminating items based on substantive grounds in conjunction with examination of modification indices, standardized residuals, chi-square statistics in light of the number of degrees of freedom, and improvements in fit statistics. Upon developing unidimensional scales through CFA, reliability and validity (convergent and discriminant) were assessed.



Enlarge 200%
Enlarge 400%

TABLE 1

The purified LSQ scale, a second-order construct comprised of nine sub-scales, was then further validated within each of the eight market segment data bases through CFA. Modifications were not made at this point. Results indicate that the LSQ scale may be applied across multiple market segments. The next section provides a detailed explanation of scale purification measures.

PURIFICATION

The primary approach for scale purification when a theoretical foundation drives survey development is to rely on CFA to ensure scale unidimensionality, followed by scale reliability and construct validity assessments. Although exploratory factor analysis (EFA) and item-to-total correlations (ITC) prior to CFA are helpful and often are used and/or recommended,3' Gerbing and Anderson argue that these procedures are most appropriate in the early stages of scale development, when the researcher is uncertain as to which items best belong together or how many dimensions underlie the data set of all items. They specifically demonstrate how developing scales through maximization of ITC and/or reliability prior to assessing construct unidimensionality can generate misleading results." Therefore, in this study, the primary means of scale purification consisted of CFA conducted with LISREL 8 to establish unidimensional scales, 39 followed by reliability assessments using Cronbach's alpha and ITC. .

Following basic descriptive analyses of returned surveys, including examination for incorrect coding, item normality, skewness, kurtosis, means, standard deviations, non-response bias and outliers, the items were grouped into apriori conceptualizations (based upon the previous literature and the focus groups) of appropriate sub-scales. Eight sub-scales were analyzed for unidimensionality using CFA in LISREL 8. The measurement model proposed for these analyses conceptualized LSQ as a second-order construct comprised of information quality, ordering procedures, timeliness, order accuracy, order quality, order condition, order discrepancy handling, and personal contact quality.

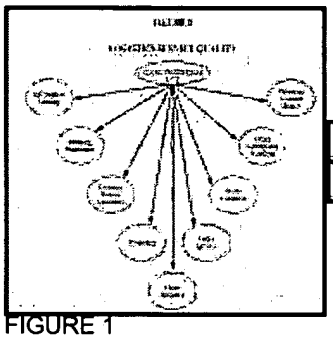
Initially, a covariance matrix was created in SPSS using the items intended to tap each of these dimensions, and it was saved as an output file to be analyzed in LISREL 8. Multiple iterations of CFA were conducted. Based on the magnitude of modification indices, standardized residuals, observed improvements in model fit indices-comparative goodness of fit index (CFI), normed fit index (NFI), chi-square with corresponding degrees of freedom-as well as substantive considerations, 27 of an initial 52 items were dropped, resulting in nine sub-scales comprised of 25 items in total (nine rather than eight, because the ordering procedures sub-scale was split into two, as explained later). This level of item deletion is not uncommon for scale development studies. When developing a scale, it is best to begin with a very large item pool. Upon completion, the final scale may contain one-fourth or even one-fifth of the original items."'

Items were deleted very selectively through eight CFA runs, each time identifying a few items for deletion. Candidates for possible deletion were initially identified through examination of modification indices. Modification indices are values provided by the LISREL program that identify items which appear to load poorly on the dimension chosen by the researcher. Through the use of these indices, the program suggests places where poorly fitting items might be better placed. For example, if the researcher specifically loads an item on a timeliness dimension but calculations indicate it might be better placed to tap order accuracy, then it may show up as a large (e.g., greater than a value of 10) modification index. These kinds of numerical "flags" were used to highlight potential problems. Final decisions to delete poor items were based on substantive grounds (see Appendix A).

Modification indices of 10 or greater were first used as indicators of potentially bad items. In subsequent iterations, acceptable modification index values became gradually more restrictive. Although the final scales possibly could have been developed in fewer runs, this would have required eliminating a larger number of items in each run. Each item deleted affects all other values. For example, deleting only one of two potentially poor items may result in the second item actually fitting the data. For this reason, a very cautious approach was taken, deleting only a few items per run, which resulted in a higher number of analysis runs. In summary, items were dropped when numerically "flagged" items were considered to be redundant or inadequate based on researcher examination. Items were "flagged" for examination by large modification indices (initially greater than 10), large standardized residuals (greater than 4), and overall poor fit statistics (CFI less than .90).

An unanticipated issue concerned the ordering procedures dimension. It immediately became apparent through examination of modification indices, standardized residuals, and substantive item analysis that the ordering procedures sub-scale consisted of two closely related scales, one on ordering procedures and the other on release

quantities. Therefore, the original scale was split into two sub-scales. This raised the LSQ scale from eight dimensions to nine, as shown in Figure 1. Once the unidimensionality of each sub-scale was demonstrated by CFA, the reliability of each was evaluated by the determination of coefficient alpha. Table 2 provides the final LSQ scale with all nine sub-scales. Alpha values are provided next to each sub-scale name where appropriate. Note that two of the nine sub-scales contain only two items. Sub-scale parsimony was a primary objective. A balance must be struck between parsimony to facilitate survey administration and construct validity, such that necessary dimensions are tapped. The goal was to find the best three items per subscale that appeared to lie at the heart of each of the nine concepts. It became apparent that two of the sub-scales were best tapped with only two items from their initial item pool. Cronbach's alpha is a meaningless calculation with a two-item scale, since its purpose is to compare each item to the remaining items in the scale as a group. Therefore, Cronbach's alpha was not calculated for the subscales of information quality and ordering procedures. For these two-item dimensions, simple correlations are reported. Item-to-total correlations (ITC) were also evaluated for all other sub-scales and are provided next to each item. All reliabilities and ITC are acceptable.



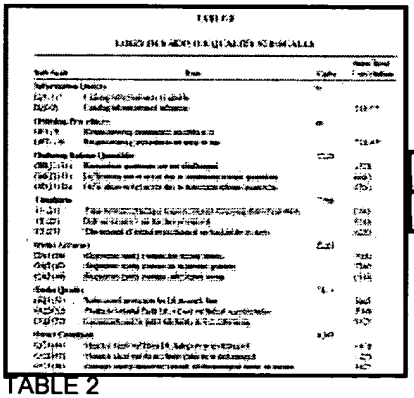
Enlarge 200%
Enlarge 400%

FIGURE 1

Table 3 presents the results of the CFA for the LSQ scale, including second-order factor loadings. A summary of item deletion decisions, including the CFA run in which each item was deleted and substantive justification for each item deleted, can be found in Appendix A.

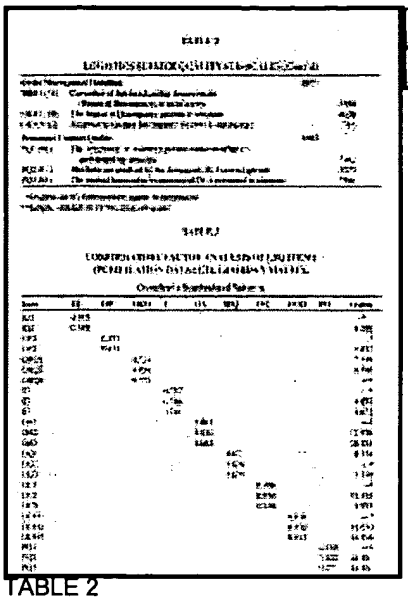
CONVERGENT AND DISCRIMINANT VALIDITY

Assessing construct validity, a term standardized by Cronbach and Meehl,42 is a complex process. An ideal form of evaluating whether items tap the constructs intended is through the multitrait-multimethod matrix.4" However, short of utilizing multiple methods to measure each construct of interest, an assessment can be gained of the validity of a construct's measurement, initially convergent and discriminant validity, through other means.



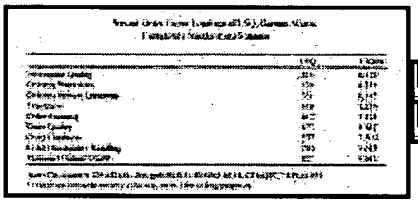
Enlarge 200%
Enlarge 400%

TABLE 2



Enlarge 200%
Enlarge 400%

TABLE 2
TABLE 3



Enlarge 200%
Enlarge 400%

Evidence of convergent and discriminant validity can be assessed in multiple ways when one has not utilized multitrait-multimethod approaches. At the basic level, if item loading values within each construct from the Lambda matrix are relatively high (i.e., greater than .50) and t values are significantly large (i.e., greater than 1.96), one has a basic comfort level for convergent validity. In actuality, because a comparison is not being made to another measure of the identical construct, this might be better termed "factorial validity."'

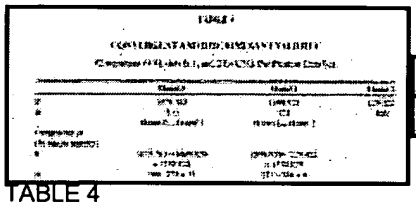
When assessing a second-order model, as is done here, a more rigorous analysis involves comparing fit statistics for multiple models, each one subsequently relaxing restrictive assumptions. This is accomplished through direct comparison of the chi-square values and degrees of freedom for three models.45 In this study, convergent validity was assessed through comparing an independence model with no traits and 25 unique factors (Model Oj with a model consisting of one trait (LSQ) and 25 unique factors (Model 1). Specifically, for Model 0, the LISREL program was run with all 25 items representing their own independent dimensions or factors (i.e., items were not loaded together on any of the LSQ dimensions). Model 1 was assessed by loading all 25 items on one latent variable-LSQ. Discriminant validity was assessed through comparison of Model I with a model proposing the structure that resulted from the CFA analysis with nine traits and 25 unique factors (Model 2). Here, Model 2 was assessed by loading appropriate items on appropriate dimensions of LSQ and then those dimensions on LSQ itself. Widaman and Bienstock, Mentzer and Bird contend that comparison of these models yields evidence of convergent (Model 0 with Model 1) and discriminant (Model 1 with Model 2) validity if the differences in chi-square values are significant.46

Table 4 demonstrates the calculations that provide evidence for convergent and discriminant validity. Model 2 provides the best overall fit in terms of chi-square. The difference in the chi-square statistics for Model 0 and Model I is significant at the a = .0001 level, thereby demonstrating evidence of convergent validity of the LSQ items. The difference in the chi-square statistics for Model 1 and Model 2 is also significant at the a = .0001 level, thereby demonstrating evidence of discriminant validity for the proposed LSQ dimensions of information quality, ordering procedures, order release quantities, timeliness, order accuracy, order quality, order condition, order discrepancy handling, and personnel contact quality.

VALIDATING LSQ ACROSS MARKET SEGMENTS

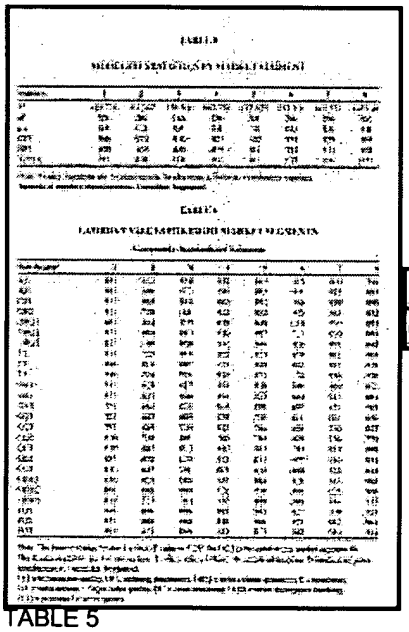
Following scale purification and tests for convergent and discriminant validity within the purification data set, the LSQ scale with nine sub-scales was further validated with the remainder of the database, which had been divided into eight distinct data sets based on market segment. For validation, the covariance matrix was created within SPSS for each of the data sets to be analyzed in LISREL 8. A single CFA run was conducted for each data set. The resulting statistics are in Tables 5, 6 and 7. The reliabilities of each sub-scale for the eight segments are presented in Table 8.

The developed LSQ scale fits the data for seven of the eight market segments (the exception is segment six, subsistence). In all seven segments, fit statistics are acceptable (refer to Table 5). CFIs are above .90, and NFIs are above .80. Chi-square values relative to degrees of freedom are best for segments one (construction - 1.8), two (electronics - 2.1), three (fuels - 2.0), and five (medical supplies - 2.1) because they are near or below a 2:1 ratio. Although chi-square values ideally should be lower for segments four (industrial supplies - 2.5) and seven (textiles - 2.3), they are acceptable. The chi-square value for segment eight (general - 5.5) is unacceptably high, due primarily to its sample size. Segment six (subsistence) has a marginally acceptable chi-square value, but unacceptable fit statistics (CFI=.874, NFI=.787). This might be due in part to the small sample size (Valid n = 157). Examination of Tables 6 and 7 reveal varying loading values by segment. This does not invalidate the scale for the segments. It merely indicates that fit is better in some segments than in others. However, close examination of Table 6 suggests various items with low loadings across multiple segments that might benefit from rewording or replacement in future research (such as, ORQ1, OQ1, OC3).



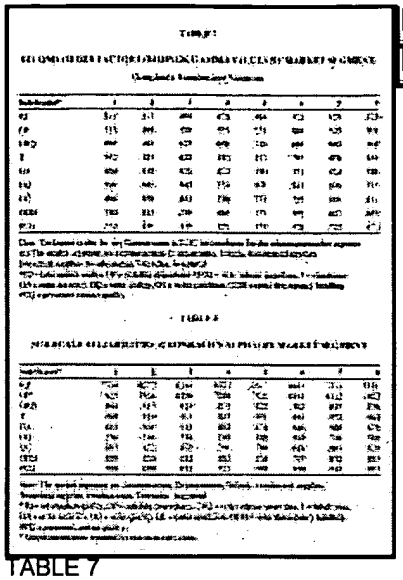
Enlarge 200%
Enlarge 400%

TABLE 4



Enlarge 200%
Enlarge 400%

TABLE 5
TABLE 6



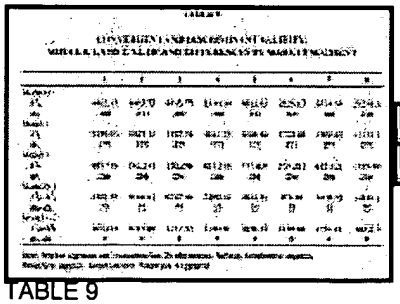
Enlarge 200%
Enlarge 400%

TABLE 7
TABLE 8

Convergent and discriminant validity were assessed in the same manner as in the purification data set, and all eight of the market segments were significant at the (a= .0001 level. Results are shown in Table 9. In each market segment, Model 2 is a better fit than Model 1 which is also a better fit than Model 0, providing evidence for convergent and discriminant validity. Thus, the developed logistics service quality scale, with its nine dimensions, withstood preliminary scrutiny with regard to validity and reliability across the purification data set and multiple market segments, providing confidence in its usefulness. Although clearly not perfect, the proposed LSQ scale offers another useful tool for quantifying customers* perceptions of logistics services.

LIMITATIONS

As with any study, this research has limitations. They concern the study sample, survey scale range, lack of scale perfection, and predictive validity.



Enlarge 200%
Enlarge 400%

TABLE 9

The sample consisted of DLA customers. Although they represent various market segments for DLA, they do not represent all types of customers of logistics services. However, one purpose of this study was to begin applying the LSQ scale development process to specific logistics organizations. Such application should continue to be pursued in future research with other organizations.

As a component of sampling, sample size varied among segments, resulting in greater variance in findings. Some of the samples were potentially too large or too small. Future research should attempt to control for this.

The survey instrument utilized a 5-point Likert "agree/disagree" scale. Some have argued that attenuation due to range restriction makes a 7-point scale more optimal.47 Therefore, findings might be limited in that respondents were not afforded opportunities for a wider range of responses. Again, resolution of this issue is left to future research.

The proposed scale is not perfect. Fit statistics could be better, fitting the data more strongly across all segments. Additionally, two dimensions of the LSQ scale were tapped with only two items. Therefore, the scale is not yet the perfect instrument for measuring logistics service quality. Future research designed to refine the proposed scale in a variety of industries should reduce this limitation. The scale developed here is an initial step toward establishing a widely usable scale.

Finally, given the scaling issues just noted, predictive validity was not assessed in this study. As an on-going effort to examine the continuum of LSQ, customer satisfaction, and share of business, the LSQ scale should be refined further, followed by tests for predictive validity (i.e., the LSQ, customer satisfaction, and share of business linkage). Again, this on-going effort is the role of future research.

MANAGERIAL IMPLICATIONS

Despite the study's limitations, managers can use the findings to improve their measurement systems. This research reveals that business customer perceptions of LSQ are multidimensional. Specifically, logisticians need to be concerned with how customers perceive information quality, ordering procedures, ordering release quantities, timeliness, order accuracy, order quality, order condition, order discrepancy handling, and personnel contact quality. In order to consistently assess and quantify customer perceptions of these service aspects, logisticians can turn to the scale developed here.

Logistics managers must quantify both their performance with respect to internal operational specifications and their performance as seen by customers with respect to the issues they find critical. The scale developed here can be used to quantify the latter. When combined with other provider attributes (such as 19

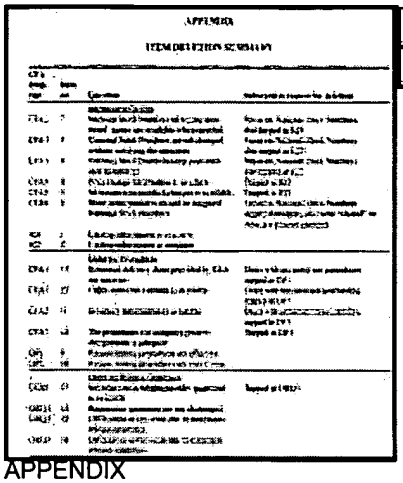
price perceptions) and other provider-relevant outcome variables (e.g., customer satisfaction, share of business awarded, word of mouth, complaint behavior, loyalty, repurchase intentions), the LSQ scale becomes an important component in determining avenues for service improvement.

CONCLUSIONS

The LSQ scale developed in this study appears to adequately fit the data collected on seven of the eight market segments of the target organization (DLA). A systematic, scientific approach established the construct validity and reliability of the LSQ scale. Future studies will need to expand the two-item scales, attempting to identify items that tap these LSQ dimensions more precisely, thus improving model fit statistics. Future research can potentially improve on the exact items within the LSQ scale dimensions, thereby improving fit statistics. Finally, the LSQ scale developed here must be evaluated across multiple industries in addition to multiple market segments within an industry.

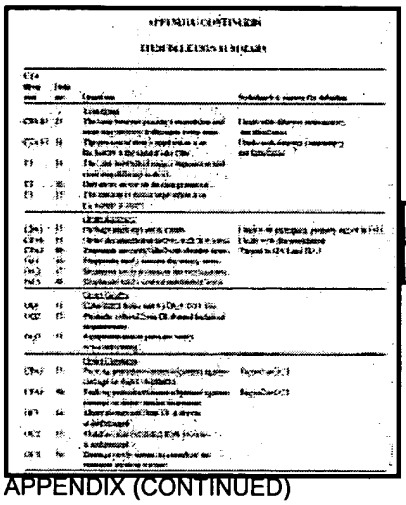
A final research implication relates to previous calls for theoretically and methodologically sound logistics research that has applicability to logistics practitioners.48 This study has attempted to follow processes called for by Mentzer and Kahn and Mentzer and Flint, namely, to build a theoretical model through antecedent literature and qualitative research, test the theoretical model via quantitative methods, and then link results to managerial action. The theory tested here was the multidimensional logistics service quality construct developed through qualitative research and refined through initial purification analysis of the split sample survey.

Although not yet perfected, this LSQ scale is another step toward enhancing logisticians' efforts to measure the impact of their operations on customers. The scale is not an operational approach that attempts to measure, from the supplier's perspective, how well providers deliver what they promised. Rather, it is a measure of how the customer perceives this delivery. Several decades of customer satisfaction research have shown that these customer perceptions have a direct impact on customer satisfaction.49 Thus, logistics customer service managers should consider using the LSQ scale in their customer questionnaires to determine customer perceptions of logistics services.



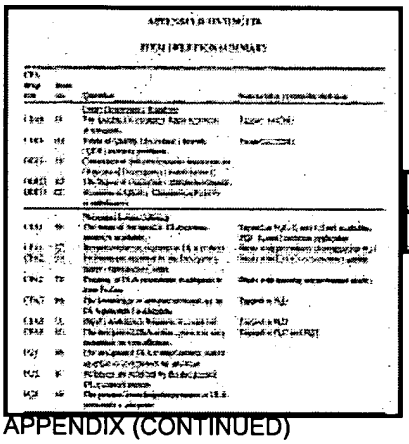
Enlarge 200%
Enlarge 400%

APPENDIX



Enlarge 200%
Enlarge 400%

APPENDIX (CONTINUED)



Enlarge 200%
Enlarge 400%

APPENDIX (CONTINUED)

[Footnote]
NOTES

[Footnote]
'Donald J. Bowersox, John T. Mentzer, and Thomas W. Speh, "Logistics Leverage," Journal of Business Strategies, 12 (Spring 1995): 36-49.
2Same reference as Note 1; and John T. Mentzer, Roger Gomes, and Robert E. Krapfel, Jr., "Physical Distribution Service: A Fundamental Marketing Concept?" Journal of the Academy of Marketin Science, 17 (Winter 1989): 53-62.

[Footnote]

3Emin Babakus, Dennis L. Pedrick, and Molly Inhofe, "Empirical Examination of a Direct Measure of Perceived Service Quality Using 'SERVQUAL' Items," in Enhancing Knowledge Development in Marketing, David W. Cravens and Peter R. Dickson, eds. (Chicago, IL: American Marketing Association, 1993), Vol. 4, pp. 8-9; Emin Babakus and Gregory W. Boller, "An Empirical

[Footnote]
Assessment of the SERVQUAL Scale," Journal of Business Research, 24 (May,1992): 253-68; Emin Babakus and Molly Inhofe, "Measuring Perceived Quality as a Multi-Attribute Attitude," AMS Proceedings, 1993, pp. 376-80; Emin Babakus and W. Glynn Mangold, "Adapting the SERVQUAL Scale to Health Care Services: An Empirical Examination," Health Service Research, 25, No. 6 (1992): 767-80; Tom J. Brown, Gilbert A. Churchill, and J. Paul Peter, "Research Note: Improving the Measurement of Service Quality," Journal of Retailing, 69 (Spring 1993): 127-39; James M. Carmen, "Consumer Perceptions of Service Quality: An Assessment of the SERVQUAL Dimensions," Journal of Retailing, 66 (Spring 1990): 33-55; J. Joseph Cronin and Steven A. Taylor, "Measuring Service Quality: A Reexamination and Extension," Journal of Marketing 56 (July 1992): 55-68; David W. Finn and Charles W. Lamb, "An Evaluation of the SERVQUAL Scales in a Retailing Setting," in Advances in Consumer Research, Rebecca H. Holman and Michael R. Solomon, eds. (Provo, UT: Association for Consumer Research, 1991), Vol. 18, pp. 483-490; Debi Mishra, Jagdip Singh, and Van Wood, "An Empirical Investigation of Two Models of Patient Satisfaction," Journal of Ambulatory Care Marketing, 4 (Winter1991): 17-36; Parasuraman, A., Valarie A. Zeithaml, and Leonard L. Berry, "A Conceptual Model of Service Quality and Its Implications for Future Research," Journal of Marketing, 49 (Fall 1985): 41-50; A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry, "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," Journal of Retailing, 63, No. 1 (1988): 12-37; A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry, "Refinement and Reassessment of the SERVQUAL Scale" Journal of Retailing, 67, No. 4 (1991): 420-50; A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry, "Research Note: More on Improving Service Quality Measurement"Journal of Retailing, 69, No.1 (1993): 140-47; A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry, "Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Future Research," Journal of Marketing, 58 (January 1994): 111-24; and Richard A. Spreng and A. K. Singh, "An Empirical Assessment of the SERVQUAL Scale and the Relationship Between Service Quality and Satisfaction," in Enhancing Knowledge Development in Marketing, David W. Cravens and Peter R. Dickson, eds. (Chicago, IL: American Marketing Association,1993), Vol. 4, pp. 1-6.

[Footnote]
4Carol C. Bienstock, John T. Mentzer, and Monroe Murphy Bird, "Measuring Physical Distribution Service Quality," Journal of the Academy of Marketing Science, 25 (Winter 1997): 31-44; and Donald P. Brensinger and Douglas M. Lambert, "Can the SERVQUAL Scale Be Generalized to Business-to-Business Services?" in Enhancing Knowledge Development in Marketing, David W. Cravens and Peter R. Dickson, eds. (Chicago, IL: American Marketing Association, 1990), p. 289. Same reference as Note 4 to Bienstock, Mentzer, and Bird.

[Footnote]
6John T. Mentzer and Kenneth B. Kahn, "A Framework of Logistics Research," Journal of Business Logistics, 16, No. 1 (1995): 231-250. 7Same references as Note 4.
Same reference as Note 6; and John T. Mentzer and Daniel J. Flint, "Validity in Logistics Research," Journal of Business Logistics, 18, No.1 (1997):199-216.

[Footnote]
9Same reference as Note 6; and William D. Perreault and Frederick A. Russ, "Physical Distribution Service: ANeglected Aspect of Marketing Management," MSUBusiness Topics, 22 (Summer 1974): 37-45.
"John J. Coyle, Edward J. Bardi, and C. John Langley, Jr., The Management of Business Logistics, 5th ed. (St. Paul, MN: West Publishing Company, 1992), pp. 38-39; Roy D. Shapiro and James L. Heskett, Logistics Strategy: Cases and Concepts, (St. Paul, Minnesota: West Publishing, 1985); and James R. Stock and Douglas M. Lambert, Strategic Logistics Management, 2nd ed. (Homewood, IL: Irwin Publishing, 1987), pp. 172-173. "Same reference as Note 2 to Mentzer, Gomes, and Krapfel. '=Same reference as Note 11.

[Footnote]
"Kenneth B. Ackerman, "Value-Added Warehousing Cuts Inventory Costs," Transportation & Distribution, 30 (July 1989): 32-35; John T. Mentzer, "Managing Channel Relations in the 21 st Century," Journal of Business Logistics, 14, No. 1 (1993): 27-42; John T. Mentzer and John Firman, "Logistics Control Systems in the 21st Century," Journal of Business Logistics, 15, No. 1 (1994): 215-228; and Clyde E. Witt, "Adding Value Accuracy Equals Quality," Material Handling Engineering 46 (April 1991): 43.
"Kenneth B. Ackerman, "Debuzzing `Value-Added,"' Transportation & Distribution, 32 (September 1991): 60.

[Footnote]
'5Bernard J. LaLonde, and Paul H. Zinszer, Customer Service: Meaning and Measurement, (Chicago, IL: National Council of Physical Distribution Management, 1976), pp. 156-159.
16Robert A. Novak, John C. Langley, Jr., and Lloyd M. Rinehart, Creating Logistics Value: Themes for the Future, Oak Brook, IL: Council of Logistics Management, 1995). "Same reference as Note 11.

[Footnote]
'8Lloyd M. Rinehart, M. Bixby Cooper, and George D. Wagenheim, "Furthering the Integration of Marketing and Logistics Through Customer Service in the Channel," Journal of the Academy of Marketing Science, 17 (Winter 1989): 63-71.
'9A. Parasuraman, Valarie A. Zeithaml, and Leonard L. Berry, "A Conceptual Model of Service Quality and Its Implications for Future Research," Journal of Marketing, 49 (Fall 1985): 41-50. 2Douglas M. Lambert, James R. Stock, and Jay U. Sterling, "A Gap Analysis of Buyer and Seller Perceptions of the Importance of Marketing Mix Attributes," AMA Educators 'Proceedings, (Chicago, IL: American Marketing Association, 1990), pp. 208.

[Footnote]
21Same reference as Note 19.
22Same refernece as Note 3.
23Philip B. Crosby, Quality Is Free: The Art of Making Quality Certain, New York: New American Library, 1979.

[Footnote]
24Tom J. Brown, Gilbert A. Churchill, and J. Paul Peter, "Research Note: Improving the Measurement of Service Quality," Journal of Retailing, 69 (Spring 1993): 127-39. 5Same reference as Note 4 to Brensinger and Lambert. :6Same reference as Note 5.
"Christopher H. Lovelock, "Classifying Services to Gain Strategic Marketing Insights," Journal of Marketing, 47 (Summer 1983): 9-20.
2"Christian Gronroos, "A Service Quality Model and Its Marketing Implications," European Journal of Marketing, 18, No. 4 (1984): 36-44.

[Footnote]
29Same refernece as Note 19.
30Same reference as Note 5.
31Same reference as Note 19.
32Same reference as Note 1.
33Same reference as Note 19.
34Same reference as Note 5.

[Footnote]
3"Bobby J. Calder, "Focus Groups and the Nature of Qualitative Marketing Research," Journal of Marketing Research, 14 (August 1977): 353-364.
6David W. Gerbing, and James C. Anderson, "An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment," Journal of Marketing Research, 25 (May 1988): 186-192.

[Footnote]
3"Same reference as Note 5 to Bienstock, Mentzer, and Bird; Robert F. DeVallis, Scale Development: Theory and Applications, (Newbury Park, CA: Sage, 1991); Rosemary P. Ramsey and Ravipreet S. Sohi, "Listening to Your Customers: The Impact of Perceived Salesperson Listening Behavior on Relationship Outcomes," Journal of the Academy of Marketing Science, 25 (Spring 1977): 127-137; and Charles H. Schwepker, Jr., O. C. Ferrell and Thomas N. Ingram, "The Influence of Ethical Climate and Ethical Conflict on Role Stress in the Sales Force," Journal of the Academy of Marketing Science, 25 (Spring 1997): 99-108. 3"Same reference as Note 36.

[Footnote]
39 Joreskog, Karl and Dag Sorbom, LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language, (Chicago, IL: Scientific Software International, Inc., 1993). "*Same reference as Note 37 to DeVallis. "Same reference as Note 37 to De Vallis.
42Lee J. Cronbach and Paul E. Meehl, "Construct Validity in Psychological Tests," Psychological Bulletin, 52 (July, 1955): 281-302.

[Footnote]
43D. T. Campbell and D. W. Fiske, "Convergent and Discriminant Validation by the MultitraitMultimethod Matrix," Psychological Bulletin, 56 (March, 1959): 81-105. 'Same reference as Note 40.
"Same reference as Note 5 to Bienstock, Mentzer, and Bird, "Measuring Physical Distribution Service Quality," Journal of the Academy of Marketing Science, 25, (Winter 1997): 31-44; and Keith F. Widaman, "Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data," Applied Psychological Measurement, 9, No. 1 (1985): 1-26. "*Same reference as Note 45.

[Footnote]
"Robert W. Lissitz and Samuel B. Green, "Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach," Journal of Applied Psychology, 60, No. 1 (1975): 10-13. 4"Same reference as Note 8.
'9Youjae Yi, "A Critical Review of Consumer Satisfaction," in Review of Marketing, Valarie A. Zeithaml, ed. (Chicago, IL: American Marketing Association, 1990), pp. 68-123.

[Author Affiliation]

John T. Mentzer
The University of Tennessee
Daniel J. Flint
Florida State University
and John L. Kent
Southwest Missouri State University

[Author Affiliation]
ABOUT THE AUTHORS

[Author Affiliation]
Tom Mentzer holds the Harry J. and Vivienne R. Bruce Excellence Chair of Business Policy in the Department of Marketing, Logistics and Transportation at the University of Tennessee. He has published more than 120 articles and papers in the Journal of Business Logistics, Journal of Marketing, Journal of Business Research, International Journal of Physical Distribution and Logistics Management, Transportation and Logistics Review, Transportation Journal, Journal of the Academy of Marketing Science, Columbia Journal of World Business, Industrial Marketing Management, Research in Marketing, and other journals.
Daniel J. Flint is Assistant Professor, Department of Marketing, Florida State University. He has published in the Journal of Business Logistics, Industrial Marketing Management, other journals, and a number of conference proceedings. His research interests include customer value and satisfaction, the marketing and leveraging of logistics services, and research methods.
John L. Kent is Assistant Professor of Logistics and Transportation, Southwest Missouri State University. He has published in the Journal of Business Logistics, International Journal of Physical Distribution and Logistics Management, Defense Transportation Journal, and several conference proceedings.

---

^ Back to Top « Back to Results < Previous Article 6 of 15 Next > Publisher Information
Print Email Mark Article Abstract, Full Text, Text+Graphics, Page Image - PDF